Sentiment Analysis: Approaches and Methods

[1]Amardeep Kaur, [2]Jagroop Kaur

[1]Punjabi University, Patiala, Punjab, India
[2]Punjabi University, Patiala,Punjab, India

Abstract: Sentiment Analysis is application that is changing the ecommerce and many other businesses around the world. It is mainly an application related with text mining and works with the integration of machine learning algorithms(ML) and deep learning algorithms. It is used to increase the business productivity and also to better the customer experience by providing meaningful data out of unstructured data. This paper explains different ways and levels to do sentiment analysis and also explains Natural Language Processing(NLP) and its different approaches. Therefore, this paper brings an overview on sentiment analysis and different techniques and approaches integrated with it.

Keywords: Lexicon, Machine Learning, NLP,Semantic Analysis, Keyword Spotting

I. INTRODUCTION

Technology is driving the industries and businesses worldwide, there has been new advancements in computer science and its technologies coming at rapid pace. Last decade has seen some revolutionary technologies like Cloud Computing, Internet of Things(IOT), Artificial Intelligence, Big Data, Software Defined Networking(SDN), Network Function Virtualization(NFV) etc are changing the world we are living. Natural Language Processing(NLP) is also one of the driving forces for artificial intelligence and text mining. NLP has now become much more powerful in recent years with the arrival of machine learning and deep learning technologies. Computers are programmed and are made to learn things, with which they are able to perform some particular operations as specified. But one thing which was not possible was to make computers think like humans. But with the popularity of Neural Networks and its advancements, deep learning and artificial intelligence is exploding its potential and using these machines can be engineered into learning, understanding and implementing actions on their own without any human intervention. There are so many examples where one can see these technologies like in Gaming Industry, Automotive Industry with Self Driving Cars, Social Networks like Facebook and Twitter for mining lots and lots of unstructured data and fetch meaningful data out of them, Ecommerce industry like Amazon or Flipkart where text mining can be used to fetch meaningful data using text mining of comments on product reviews and get the customer sentiment on some specific type of products. Text Semantics can be used to understand the meaning of text when the words are integrated to make sentences. These words also have a lexical and contextual relations between them that automatically leads them to various other relationships and hierarchies. Semantics is the major thing which is used to analyze the relationships and fetch meaningful data. Although Semantics does not need any structure of the text, but is totally related with the context and meaning, but there are some cases where the arrangement and syntax of words helps us in fetching the meaningful data and in differentiate things like "MI is a technology company" from "MI is a movie series of Tom Cruise".

II. SENTIMENT ANALYSIS

Sentiment Analysis is a popular text analytical application and it is used in integration with large number of applications like web sites, mobile apps etc. The major focus of sentiment analysis is to analyze sentiments from large text resource mostly in unstructured manner. Different text resources are like Survey which can be like Google Opinions, Posts or Tweets on Social Network[2][3] Pages of organizations endorsing some product, some social network pages of famous personalities like Politicians[4], Artists, Sportspersons etc. Sentiment Analysis can also be used to get the meaningful data using the reviews of customers[1] on some products which the customers are using and find out what people think of the product in negative, positive or neutral manner. The focus of sentiment analysis is to analyze the text body and to find and then to understand the opinion and sentiment expressed on the basis of factors like mood or modality etc. Sentiment analysis works best with the subjective context. The reason for more emphasis on subjective context over objective context is because with objective context, text is mainly made of normal sentences with no expression of emotions or feelings. Textual Data is mostly unstructured and is of two types i.e. Subjective,

which is also known as Opinion Based and Objective, which is factual based. Sentiment Analysis uses NLP techniques, lexical resources, machine learning and deep learning techniques to gather data like emotions and feelings etc. and then use the data gathered to compute the polarity of the document. Polarity of the document can be described in either the document expresses negative, positive or neutral sentiment. There is also much more advanced analysis with complex emotions like sadness, anger, or sarcasm etc. Sentiment Analysis is also based on to different levels:

Document – This level fetches the sentiment at the document level. In this level, document is based on single topic.

Sentence – This level fetches the sentiment at the sentence level. It looks for a sentiment in every sentence. Sentence which does not have any opinion means that it has a neutral view. It is also integrated with subjectivity classification. Subjective Statement displays the polarity in the statement and expresses if the sentence is good or bad. It is very easy to get a sentiment from it.

Entity or Aspect based – It is the detailed based analysis. It looks for the aspect of the text. For example, in a movie review by the customer, "Acting was good, but the screenplay was bad." In this example, aspects are acting and screenplay. Sentiment Analysis integrates two different tasks in one entity level. First, it finds the aspects in the text and secondly, it classifies the text according to the aspect. Therefore it produces much better results than the Document or Sentence based levels.

III.   METHODOLOGIES

In largest number of cases, text analysis is done at the document level. During polarity analysis, scores are assigned to the levels of expressions i.e. scores are assigned to find the levels of positivity, negativity in the emotions or feelings expressed in the document and a document is accessed as a whole on the basis of the summarized score.Various types of sentiment analysis is explained below:

Keyword Spotting – In this way, text is categorized on the basis of some unambiguous words which can be the part of the sentence or the document. These words can have some special meaning in respective to sentiment analysis.

Applications: Speech Processing, Image Processing, Text Analysis, Sentiment Analysis

Lexical Affinity – It provides a probabilistic similarity to some words for some particular emotion or feeling.

Applications: Text Analysis, Sentiment Analysis

Statistical Method – This method calculates the target of affective keywords. It also calculates the co-occurrence numbers on the basis of large training corpus.

Applications: Text Analysis, Sentiment Analysis

Sentiment Analysis is categorized into three major categories i.e. Lexicon based, Machine Learning Based and the integration both, which is also known as Hybrid Approach. Machine Learning using the mixture of ML algorithms and linguistic features. Lexicon method is dependent on sentiment lexicon which can be a dictionary, or a book containing words. Lexicon in my work can be a dictionary of words which is used to analyzing sentiments as all the words in the dictionary are associated with some special scores which integrated with the positioning of the words in a sentence or a document are used to analyze the sentiments in the document. Sanjida Akter and Muhammad Tareq Aziz[3] have shown how to predict the sentiment of the status post which is regarded as a unstructured dataset. Lexicon Based Approach works best in this kind of works

behind a status post of Facebook which in nature of unstructured dataset, cross language domain and noisy. Traditional opinion mining is not efficient to gather sentiments from social media giant like Facebook and lexicon based dictionary approach works efficiently in such kind of works.

Hybrid Approach is the composite type integrating both machine learning approach and lexicon based approach. Sentiment Analysis further division and sub categories of categories are defined in Figure 1.1:
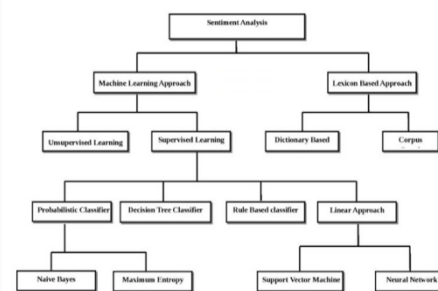

Figure 1.1 – Sentiment Analysis Technique

IV.   CONCLUSION

Sentiment Analysis is used by different industries to expand their business growth. It uses the integration of machine learning and deep learning algorithms and perform text mining to fetch meaningful data out of unstructured data. The data fetched brings lots of benefits to the customers and businesses and let them select the right product or service by providing them better insight and emotions of customers in the form of reviews, comments, ratings etc. Social Network Sites are used with almost all the fields for marketing. Sports, Movies, Music, Politics, Products etc uses Social Network heavily with lots of activities on Twitter[6-10] or Facebook etc Social Network Sites. Sentiment Analysis can be done on social sites like on Facebook Pages of some

famous personalities, Political Parties, Product etc pages and using Facebook Open API and then by collecting the data fetched from Facebook Pages and analyze the comments, likes, reactions and reviews on Facebook Posts. This paper is an overview on what sentiment analysis is and its various methods and approaches along with the industries in which it can be utilized.

REFERENCES

[1] Anna Baj-Rogowska(2017),” Sentiment Analysis of Facebook Posts:the Uber case”, IEEE International Conference on Intelligent Computing and Information Systems (ICICIS).

[2] Antonio Teixeira and Raul M.S. Laureano(2013),”Data extraction and preparation to perform the sentiment analysis using open source tools”.

[3] Sanjida Akter and Muhammad Tareq Aziz(2016),” Sentiment Analysis On Facebook Group Using Lexicon Based Approach”,IEEE.

[4] Saud Alashri, Srinivasa Srivatsav Kandala, Vikash Bajaj, Roopek Ravi, Kendra L. Smith and Kevin C. Desouza(2016),” An Analysis of Sentiments on Facebook during the
2016 U.S. Presidential Election”, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).

[5] N. AZMINA M. ZAMANI, SITI Z. Z. ABIDIN, NASIROH OMAR, M. Z. Z.ABIDEN(2013),”Sentiment Analysis: Determining People's Emotions in Facebook”, Universiti Teknologi MARA, Malaysia.

[6]Vishal A. Kharde and S.S. Sonawane(2016),” Sentiment Analysis of Twitter Data: A Survey of Techniques”, International Journal of Computer Applications (0975 – 8887) Volume 139 – No.11.

[7] Shufeng Xiong, Hailian Lv , Weiting Zhao, Donghong Ji(2017),” Towards Twitter sentiment classification by multi-level sentiment-enriched word embeddings”, Elsevier.

[8] Nehal Mamgain, Ekta Mehta, Ankush Mittal and Gaurav Bhatt(2016),” Sentiment Analysis of Top Colleges in India Using Twitter Data”, International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT).

[9] Aliaksei Severyn abd Aliaksei Severyn(2015),” Twitter Sentiment Analysis with Deep Convolutional Neural Networks”, ACM. ISBN 978-1-4503-3621-5/15/08.

[10] Jˆonatas Wehrmann, Willian Becker, Henry E. L. Cagnini, and Rodrigo C. Barros(2017),” A Character-based Convolutional Neural Network for Language-Agnostic Twitter Sentiment Analysis”,IEEE.

[11] Nhan Cach Dang, Fernando De la Prieta, Juan Manuel Corchado and María N. Moreno(2016),” Framework for Retrieving Relevant Contents Related to Fashion from Online Social Network Data, Springer.